Victoria Hsiao

Caltech Amgen Scholars Program – Progress Report #2
iGEM 2008 Smolke Lab
August 1st 2008

Making Folate Factories in *E. coli*

Progress to Date

This past month, I have spent most of my time cloning the four genes I extracted from the *Lactoccocus lactis* genome (*folB, folKE, pabA, pabB*) into constructs containing other parts from the Biobricks ™ registry. Cloning is a process which entails digesting, ligating, transforming, screening, inoculating, and then sequencing each gene. As described in my last progress report, the goal of cloning is to insert each gene into a Biobricks ™ plasmid with the following Biobricks ™ parts: a promoter, a ribosomal binding site, the gene, and then a double terminator. At the time of my last progress report, I had only just started on the first cloning step, which entailed adding a ribosomal binding site (RBS) to each individual gene cloned from the genome. Since then I have taken the products from the first cloning step and added a constitutive promoter to each individual gene. Currently, I am working on cloning the products from the second step (promoter + RBS + gene) in front of a double terminator in the plasmid psB2K3. PsB2K3 is an inducible copy plasmid, which means that replication of the plasmid can be switched from low copy to high copy via the addition of some small biological reagent to the liquid culture; In the case of psB2K3, this reagent is isopropyl β-D-1-thiogalactopyranoside (IPTG) .

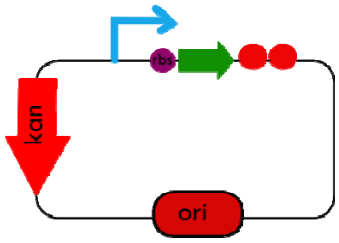

Figure 1: Complete plasmid construct. The blue arrow represents the promoter, the purple circle is the ribosomal binding site (RBS), the green arrow is the gene, the red circles are the double terminator, and the rest of the figure is the plasmid with kanamycin resistance and an origin of replication.

This addition of the double terminator to each of my constructs is the third and final step for creating fully functional plasmids for each individual gene. With these finished constructs I will be able to begin testing the effects of overexpression of each of these genes individually on total folate production in *E.coli*. In addition, I also need to clone two more constructs containing *folB + folKE* and *pabA + pabB*, such that I can also test the effects of overexpression of multiple genes in the folate biosynthesis pathway or the para-aminobenzoic acid pathway. Currently, I have *folB + folKE* behind a promoter, which needs to be put in front of the terminator (one cloning step away) ; I also have *pabA* behind a promoter, and am trying to successfully clone *pabB* in front of the terminator such that I can ligate the two together (two cloning steps away).

Data Analysis Techniques

Currently I'm testing out and optimizing the protocol for detecting folate and para-aminobenzoic acid (pABA). For folate, I'm using a microbiological assay which will entail growing a

folate-dependent strain of bacteria (*Enterococcus hirae*) in the cell lysates of my samples and then measuring the optical absorbance of the cultures to determine the relative folate concentrations. Currently, I am still working out the standard curve for the assay by using 0.0, 0.5, 1, 2, 4, 6, 8, and 10 ng of folate. Ideally the standard curve would be linear, and my results thus far have not been as linear as they could be (Fig. 2). My current theories are that I'm incubating the assay samples for too long and the folate-dependent strain is saturating, I'm adding too much folate to my samples, or there is a contaminating source. Therefore, I plan on reducing the incubation time from 18 hours to 12 hours, and taking sample data every 4 hours after inoculation for 12 hours.

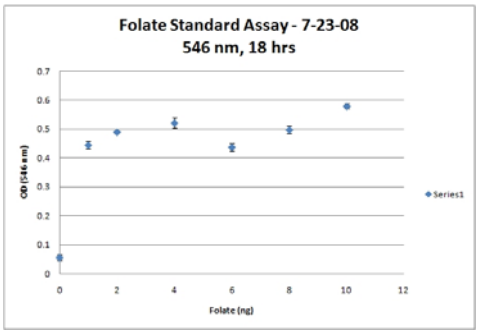

Figure 2: Best Folate Standard Assay results to date

For the detection of pABA, I plan on using high performance liquid chromatography (HPLC), which entails running samples through a chromatography column which separates based on polarity. Currently the HPLC protocol runs at 14 minutes per sample, with the pABA retention times around 6 minutes (Fig. 3).

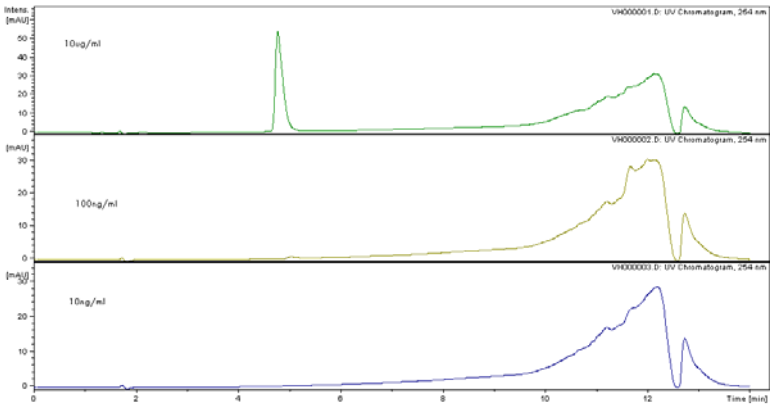

Figure 3: pABA HPLC Standard Assay Results. Note the distinct peaks at approx. 5 minutes which correspond to pABA concentration.

Problems Encountered

One of the main problems that I've encountered, and that I have still not been able to overcome, is that I've been unable to extract the folate operon, *folC* or *folP* from the *L.lactis* genome.

Many attempts have been made to overcome this obstacle, such as designing new primers, running gradient polymerase chain reactions (PCRs) to try and find the optimal annealing temperatures, and trying primers from other genes upstream of the gene in question. Although none of these efforts have been successful, we now suspect that it is because the primer design was based off of a sequence which doesn't completely match the actual sequence of the *L.lactis* strain we received in the mail from ATCC. Based on sequencing results from the genes which we were able to clone, we now believe that the strain we have is a homologous strain from the strain whose sequence I've been basing my primer design off of. As such, the primers I designed probably aren't binding to the target gene sequence. There are several strategies I can employ going forward, such as designing primers every hundred bases along the genes I have already sequenced, or using the primers from the primary literature for the entire folate operon (which were designed for yet another strain of *L.lactis*. However, I have not had the time to pursue these other options, and even if I were able to successfully clone these genes out of the genome, I would not have enough weeks left in the summer to make entire constructs.

As previously mentioned, I need to optimize the incubation times for my folate assay such that the standardization curve appears linear. Previously, we had some issues with the folate assay because we had used Triton X-100 in place of the recommended Tween 80. This was corrected after it was found that the purpose of Tween 80 was not as a surfactant; Tween 80 turned out to be important because it was derived from oleic acid and was found to minimize the lag effects of oleic acid on the growth of the folate detecting strain. I also had some issues with the pABA HPLC assay where the protocol we were using required 110 minutes per sample and did not produce good readings. We were able to find another protocol with different buffers which reduced the runtime to only 14 minutes per sample and gave good results.

Research Goals

To date, my research goals remain pretty much the same as they were at the beginning of the summer, with the exception of the two folate genes (*folP* and *folC*) that I have not been able to clone out of the genome. We were originally also going to clone the entire folate operon and compare overexpression of the entire operon with overexpression of individual genes, but that was quickly discarded early on when we were also unable to clone that out of the genome. We suspect that this also because of differences in sequence between the actual strain and the strain I used for primer design, since *folC* is the last gene in the operon, and what I based the reverse primer off of.

The goal is still to metabolically engineer *E. coli* to overproduce folate by testing different combinations of gene overexpression from both the folate synthesis pathway and the pABA synthesis pathway. Once I finish the third and final cloning step, I will be able to start on these experiments.

Interaction with Mentor/Co-mentors

We have team meetings every week with our two graduate student co-mentors and our mentor, which are productive because it keeps everyone up to date on your individual project. Usually if I have questions I ask one of my co-mentors.